Proyecto 1

Inteligencia De Negocios

Jairo Nicolás Gómez

Paola Campiño

Felipe Duque

Análisis de Sentimientos en Películas

Abril 2 de 2023

Tabla de contenido

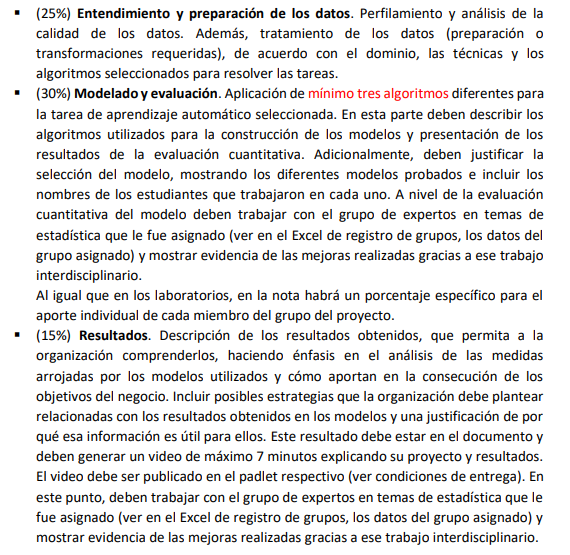
# Contenido

[1. Contenido 2](#_Toc130984028)

[2. Entendimiento del negocio y enfoque analítico 3](#_Toc130984029)

[3. Modelado y evaluación: 3](#_Toc130984030)

[4. Resultados: 3](#_Toc130984031)



# Entendimiento del negocio y enfoque analítico

|  |  |
| --- | --- |
| Oportunidad/problema Negocio | La oportunidad que vemos en este proyecto es poder ver y analizar las opiniones de los espectadores para mejorar los factores que son más criticados en las películas y mantener los aspectos más positivos y que son mejor resaltados por los espectadores. |
| Enfoque analítico (Descripción del requerimiento desde el punto de vista de aprendizaje automático) | Llegar a obtener un modelo con una alta precisión para analizar las reseñas sobre las películas y lograr clasificar de la mejor manera si son negativas o positivas. |
| Organización y rol dentro de ella que se beneficia con la oportunidad definida |  |
| Técnicas y algoritmos a utilizar | Utilizaremos una técnica de clasificación y los algoritmos de regresión logística, arboles de decisión y SVM (máquinas de vectores de soporte). |

# Modelado y evaluación:

Como ya se ha mencionado antes se ha usado tres técnicas y algoritmos para realizar los modelos. A continuación se presenta una prequeña explicación de los resultados:

# Árbol de decisión:

Para la implementación del árbol de decisión por este algoritmo se planteó 3 diferentes modelos. El primero con un 0.691 en la que se define el número mínimo de elementos para generar un nodo de 2 elementos, usando como métrica la entropía y una altura por default.

precision recall f1-score support

0 0.68 0.73 0.70 498

1 0.71 0.65 0.68 502

accuracy 0.69 1000

macro avg 0.69 0.69 0.69 1000

weighted avg 0.69 0.69 0.69 1000

Para el segundo modelo se intentó encontrar el mejor árbol posible del que salío que los parámetros deberían ser que la métrica sería entropía, la altura sería de 28 y que el numero mínimo de elementos debía estar 4 del que si se llegó a las siguientes métricas:

precision recall f1-score support

0 0.66 0.80 0.72 498

1 0.75 0.59 0.66 502

accuracy 0.69 1000

macro avg 0.70 0.69 0.69 1000

weighted avg 0.70 0.69 0.69 1000

# Resultados: